ELSEVIER Linear Algebra and its Applications 321 (2000) 153–173

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

The asymptotic covariance matrix of maximum-likelihood estimates in factor analysis: the case of nearly singular matrix of estimates of unique variances[☆]

Kentaro Hayashi [a], Peter M. Bentler [b,*]

[a] *Department of Psychology, Utah State University, Logan, UT 84322-2810, USA*
[b] *Departments of Psychology and Statistics, University of California, P.O. Box 951563, Los Angeles, CA 90095-1563, USA*

Received 26 March 1998; accepted 3 July 2000

Submitted by H.J. Werner

Abstract

This paper is concerned with the asymptotic covariance matrix (ACM) of maximum-likelihood estimates (MLEs) of factor loadings and unique variances when one element of MLEs of unique variances is nearly zero, i.e., the matrix of MLEs of unique variances is nearly singular. In this situation, standard formulas break down. We give explicit formulas for the ACM of MLEs of factor loadings and unique variances that could be used even when an element of MLEs of unique variances is very close to zero. We also discuss an alternative approach using the augmented information matrix under a nearly singular matrix of MLEs of unique variances and derive the partial derivatives of the alternative constraint functions with respect to the elements of factor loadings and unique variances.

AMS classification: 62H25; 62F12

Keywords: Augmented information matrix; Factor loadings; Heywood case; Standard errors

☆ This work was supported by National Institute on Drug Abuse grant DA01070.

* Corresponding author. Tel.: +1-310-825-2893; fax: +1-310-206-4315.
E-mail addresses: khayashi@coe.usu.edu (K. Hayashi), bentler@ucla.edu (P.M. Bentler).

PII: S 0 0 2 4 - 3 7 9 5 (0 0) 0 0 2 2 2 - 6

1. Introduction

This paper is concerned with the asymptotic covariance matrix (ACM) of maximum-likelihood estimates (MLEs) of factor loadings and unique variances when one element of MLEs of unique variances is nearly zero, i.e., when the MLE of the matrix of unique variances is nearly singular. It has been known for a long time that this situation occurs quite frequently in practice, e.g., Jöreskog [14] noted that "...improper solutions are quite frequent. Out of the 11 sets of data considered only two sets (Data 1 and 5) have proper solutions for all values of k_0 (the number of factors). This is a most remarkable result" (p. 473). While he developed an effective computational method to yield parameter estimates in spite of this problem, he did not provide standard error formulas that could be applied in this circumstance.

The formulas for the ACM of MLEs of factor loadings and unique variances for the regular case were obtained by Lawley [16], and they were systematically presented in [18]. There are some mistakes in the formulas presented by Lawley and Maxwell [18] and the mistakes were corrected by Jennrich and Thayer [13]. Jennrich [9] and Lawley [17] introduced the augmented information matrix approach which simplifies the process of obtaining the ACM of MLEs of factor loadings and unique variances.

The formulas presented in [13,18] involve a reciprocal of unique variances. When one of the MLEs of unique variances approaches zero, the reciprocal diverges to plus infinity. As a result, the estimated ACM computed using the standard formulas in [13,18] breaks down as one element of the MLEs of unique variances gets very close to zero. Thus, it is desirable to come up with alternative formulas for the ACM of MLEs of factor loadings and unique variances which do not involve a reciprocal of unique variances. Furthermore, the standard formulas are likely to run into computational instabilities if some of the unique variances are very small, and rounding errors can result in disastrous outcomes of the computations. The alternative methods, on the other hand, should be more stable, avoiding problems with rounding errors. The purpose of this paper is to give such alternative formulas and also an alternative procedure to obtain the ACM that can be used even when the MLE of the matrix of unique variances is nearly singular.

Jennrich and Clarkson [11] developed a method to compute approximate standard errors which makes use of a jackknife-like procedure. Their paper dealt with the Heywood case (i.e., case with a zero value of MLE of a unique variance) problem by expressing elements of the differentials of factor loadings without the inverse of unique variances. In this paper, we present exact (both standard and alternative) formulas for the ACM of MLEs of factor loadings and unique variances by making use of the differentials obtained by Jennrich and Clarkson [11].

In addition to the formulas mentioned above, we also provide another approach to compute the ACM of MLEs of factor loadings and unique variances using the augmented information matrix under a nearly singular matrix of MLEs of unique variances. The partial derivatives of the constraint functions given by Jennrich [9] involve reciprocals of elements of unique variances, so use of the standard partial derivatives creates a problem as one of the elements of MLEs of unique variances approaches zero. We use an alternative constraint function which does not involve reciprocals of unique variances, and derive the partial derivatives of the alternative constraint functions with respect to the elements of factor loadings and unique variances to implement the augmented information matrix approach.

We first introduce the formulas for MLEs of factor loadings and unique variances and the ACM for the regular case in Section 2. An alternative formula for MLEs of factor loadings is reviewed in Section 3, along with a discussion of several issues associated with the case of a nearly singular matrix of MLEs of unique variances. Then, in Section 4, we present our alternative formulas for the ACM of MLEs of factor loadings and unique variances. Section 5 presents both the standard and alternative formulas derived from the differentials given by Jennrich and Clarkson [11]. The augmented information matrix approach with an alternative constraint function and the partial derivatives follows in Section 6. Section 7 is a brief conclusion.

2. The ACM of MLEs of factor loadings and unique variances: Ψ version

Let $x_i,\ i = 1, \ldots, n$, be a $p \times 1$ random vector of observations with the mean vector 0 and the covariance matrix Σ, Λ be a $p \times m$ matrix of factor loadings, f_i be an $m \times 1$ vector of the common factors, and ϵ_i be a $p \times 1$ vector of the unique factors. The factor analysis model is given by $x_i = \Lambda f_i + \epsilon_i$ with $\boldsymbol{E}(f_i) = 0$, $\mathrm{Cov}(f_i) = I_m$, $\boldsymbol{E}(\epsilon_i) = 0$, $\mathrm{Cov}(\epsilon_i) = \Psi$, $\mathrm{Cov}(f_i, \epsilon_i) = 0$, where Ψ is a positive definite (or positive semidefinite) diagonal matrix. Then the covariance matrix Σ is expressed as $\Sigma = \Lambda\Lambda' + \Psi$. In MLE, we further assume that x_i's are random samples from a multivariate normal population with mean vector 0 and covariance matrix Σ, and the constraint that $\Lambda'\Psi^{-1}\Lambda$ is diagonal is imposed for Λ to be identified. It is known that the MLEs of Λ and Ψ are obtained by solving the following two equations:

$$\Lambda = \Psi^{1/2}\Omega^*(\Theta^* - I_m)^{1/2}, \tag{1}$$

$$\Psi = \mathrm{diag}(S - \Lambda\Lambda'), \tag{2}$$

where Θ^* is an $m \times m$ diagonal matrix whose elements are the first m largest eigenvalues of $\Psi^{-1/2}S\Psi^{-1/2}$, Ω^* a $p \times m$ matrix whose columns are normalized (i.e., $\Omega^{*\prime}\Omega^* = I_m$) eigenvectors corresponding to the first m largest eigenvalues of $\Psi^{-1/2}S\Psi^{-1/2}$, $\mathrm{diag}(A)$ denotes the diagonal matrix whose elements are the diagonal elements of the square matrix A, and S is the sample covariance matrix.

Let $\widehat{\Lambda}$ and $\widehat{\Psi}$ be the MLEs of Λ and Ψ, $\widehat{\lambda} = \mathrm{vec}(\widehat{\Lambda})$ and $\widehat{\psi} = \mathrm{vdg}(\widehat{\Psi})$, where $\mathrm{vec}(\widehat{\Lambda})$ denotes the pm-vector listing m columns of the $p \times m$ matrix $\widehat{\Lambda}$ starting from the first column, and $\mathrm{vdg}(\widehat{\Psi})$ denotes the diagonal elements of $\widehat{\Psi}$ arranged as a p-vector. Anderson and Rubin [3] established the asymptotic multinormality of $\sqrt{n}(\widehat{\lambda} - \lambda)$ and $\sqrt{n}(\widehat{\psi} - \psi)$ under the following three assumptions: (i) $\Phi \odot \Phi$ is nonsingular, i.e., the determinant $|\Phi \odot \Phi| \neq 0$, where Φ is defined in (13), and $\odot$ is the Hadamand product; (ii) Λ and Ψ are identified by the constraint that $\Lambda'\Psi^{-1}\Lambda$ is diagonal and the diagonal elements are different and ordered; (iii) the sample covariance matrix S converges to Σ, in probability, and $\sqrt{n}(S-\Sigma)$ has an asymptotic multinormal distribution. (Actually, the joint asymptotic multinormality of $\sqrt{n}((\widehat{\lambda}'\widehat{\psi}')'-(\lambda'\psi')')$ also holds. See, e.g., [2].)

We now present the standard formulas for the ACM of MLEs of factor loadings and unique variances. The original elementwise formulas were given by Lawley and Maxwell [18], with correction by Jennrich and Thayer [13]. The formulas that we give are a matrix version of the identical results, which is slightly modified from the matrix results given by Hayashi and Sen [7]. (Note: The proof for the equivalence between the elementwise formulas and the matrix formulas for A and B are given in Appendix A.3. The matrix formulas for E and Φ were given by Lawley and Maxwell [18].) The formulas for the ACM are:

$$\begin{pmatrix} \mathrm{Var}(\widehat{\lambda}) & \mathrm{Cov}(\widehat{\lambda},\widehat{\psi}) \\ \mathrm{Cov}(\widehat{\psi},\widehat{\lambda}) & \mathrm{Var}(\widehat{\psi}) \end{pmatrix} = \left(\frac{1}{n}\right)\begin{pmatrix} A+2B'EB & 2B'E \\ 2EB & 2E \end{pmatrix}, \tag{3}$$

where A (of order $pm\times pm$), B (of order $p\times pm$), and E (of order $p\times p$) are as follows:

$$A=\{M\otimes\Sigma+(M\otimes\Lambda M)(\mathrm{diag}(K_m\gamma^*))(I_m\otimes\Lambda')\}-\{A_1\odot A_1'\odot A_2\}, \tag{4}$$

$$\begin{aligned} B = &-\{\Psi^{-2}\Lambda(\Theta-I_m)^{-1}\otimes 1_p'\} \\ &\odot\{1_m'\otimes\Psi+(1_m'\Theta\otimes\Lambda)(\mathrm{diag}(K_m\theta^b))(I_m\otimes\Lambda')\}, \end{aligned} \tag{5}$$

$$E=(\Phi\odot\Phi)^{-1}, \tag{6}$$

with

$$A_1=1_m\otimes\Lambda\otimes 1_p'-(\mathrm{diag}(\lambda))(I_m\otimes 1_p1_p'), \tag{7}$$

$$A_2=\Theta\Theta^{\ddagger}\Theta\otimes 1_p1_p', \tag{8}$$

$$\gamma^*=\mathrm{vec}((\Theta-I_m)^2\Theta^{\ddagger}-1_m1_m'+\left(\tfrac{1}{2}\right)I_m), \tag{9}$$

$$M=\Theta(\Theta-I_m)^{-1}, \tag{10}$$

$$\Theta^{\ddagger}=\mathrm{vec}^{-1}(((I_m\otimes\Theta-\Theta\otimes I_m)^2)^{+}1_{m^2}), \tag{11}$$

$$\theta^b=(I_m\otimes\Theta-\Theta\otimes I_m-2\,\mathrm{diag}(\mathrm{vec}(\Theta(\Theta-I_m))))^{-1}1_{m^2}, \tag{12}$$

$$\Phi=\Psi^{-1}-\Psi^{-1}\Lambda(\Theta-I_m)^{-1}\Lambda'\Psi^{-1}, \tag{13}$$

where Θ is an $m\times m$ diagonal matrix whose elements are the first m largest eigenvalues of $\Psi^{-1/2}\Sigma\Psi^{-1/2}$, vec^{-1} the inverse operation of vec, i.e., $\mathrm{vec}(\Theta^{\ddagger})=((I_m\otimes$

$\Theta - \Theta \otimes I_m)^2)^+ 1_{m^2}$ with $\Theta^{\ddagger}$ being an $m \times m$ matrix, $+$ in (11) the Moore–Penrose inverse, $\otimes$ the Kronecker product, K_m the $m^2 \times m^2$ commutation matrix defined such that $K_m \operatorname{vec}(G) = \operatorname{vec}(G')$ for any $m \times m$ matrix G, $\operatorname{diag}(z)$ denotes the diagonal matrix whose diagonal elements are vector z, I_m is the m-dimensional identity matrix, and 1_m is the m-vector whose elements are all 1's.

Note that Eqs. (5) and (13) involve inverses of unique variances. Also, the estimate of the $(1, 1)$ element of Θ, i.e., the estimate of the largest eigenvalue of $\Psi^{-1/2}\Sigma\Psi^{-1/2}$, gets very large as one of the MLEs of unique variances approaches zero. Thus, Eqs. (8)–(12) also need to be modified.

3. The case with a nearly singular matrix of MLEs of unique variances

We now discuss the case where one element of the matrix of MLEs of unique variances is very close to zero, i.e., $\widehat{\Psi}$ is nearly singular. (We deal only with exploratory factor analysis in this paper. For confirmatory factor analysis, see, e.g., [6].) In this section, we discuss several issues associated with a nearly singular matrix of MLEs of unique variances.

First, note that the assumption of positive semidefiniteness of Ψ, i.e., $\{\Psi\colon \Psi \geqslant 0\}$, is actually $\{\Psi\colon 0 \leqslant \Psi \leqslant \Sigma\}$, since $\Lambda\Lambda'$ is also positive semidefinite. We further assume that the true parameter values Ψ_0 of unique variances lie in the interior of the parameter space $\{\Psi\colon 0 \leqslant \Psi \leqslant \Sigma\}$; thus Ψ_0 is nonsingular. (One of the regularity conditions for the asymptotic normality of MLEs requires that the neighborhood around the true parameter values needs to be inside the parameter space. See, e.g., assumption (vi) of Theorem 2 of Anderson and Amemiya [2, p. 764].) The assumption of nonsingularity of Ψ_0 implies that the probability of obtaining a nonsingular $\widehat{\Psi}$ (i.e., $\widehat{\Psi}$ in the interior of $\{\Psi\colon 0 \leqslant \Psi \leqslant \Sigma\}$) approaches unity as sample size increases.

Second, the MLEs of factor loadings can be computed even when one element of MLEs of unique variances is exactly zero, by using the eigenvalues ν_r^* of $U^*\Psi U^{*\prime}$, where U^* is the Cholesky factor of S^{-1} (i.e., $S^{-1} = U^{*\prime}U^*$), instead of using the eigenvalues θ_r^* of $\Psi^{-1/2}S\Psi^{-1/2}$. We will state this as Observation 1, as follows.

Observation 1 [12, 21, 22].

(i) *Assume that S is positive definite. Let ν_r^* be the eigenvalues of $U^*\Psi U^{*\prime}$ with $S^{-1} = U^{*\prime}U^*$, where U^* is an upper triangular matrix with positive diagonal elements obtained by Cholesky decomposition of S^{-1}, and let ζ_r^* be the normalized eigenvectors corresponding to ν_r^*. Let N^* be an $m \times m$ diagonal matrix whose elements are the m smallest eigenvalues $\nu_1^*, \ldots, \nu_m^*$ ($\nu_1^* < \cdots < \nu_m^*$), and let Z^* be a $p \times m$ matrix whose* rth *column is ζ_r, $r = 1, \ldots, m$. Then we have the following relationships*:

$$\Theta^* = N^{*-1}, \tag{14}$$

$$\Omega^* = \Psi^{1/2} U^{*\prime} Z^* N^{*-1/2}. \tag{15}$$

(ii) *The matrix of factor loadings is computed without Θ^* and Ω^*, as follows*:

$$\Lambda = U^{*-1} Z^* (I_m - N^*)^{1/2}. \tag{16}$$

Thus, alternatively, the MLEs of Λ and Ψ are obtained by solving Eqs. (16) *and* (2), *instead of* (1) *and* (2).

See Appendix A.1 for the proof of Observation 1. Note that Eq. (1) is derived from $(S - \Sigma)\Psi^{-1}\Lambda = 0$, while Eq. (16) is derived from $(S - \Sigma)S^{-1}\Lambda = 0$. In fact, the three equations $(S - \Sigma)\Psi^{-1}\Lambda = 0$, $(S - \Sigma)S^{-1}\Lambda = 0$, and $(S - \Sigma)\Sigma^{-1}\Lambda = 0$ are equivalent, except that positive definiteness of Ψ, S, and Σ are assumed, respectively [19, Theorems 4.2 and 4.3]. For example, equivalence of $(S - \Sigma)\Psi^{-1}\Lambda = 0$ and $(S - \Sigma)\Sigma^{-1}\Lambda = 0$ can be shown by making use of the identity $\Psi^{-1}\Lambda\Theta^{-1} = \Sigma^{-1}\Lambda$ [18, Eq. (4.7)] to convert one to the other (see Appendix A.8).

Third, although the MLEs of factor loadings can be computed even when one element of MLEs of unique variances is exactly zero, caution has to be taken in interpreting a zero value of $\widehat{\psi}_j$ in the same way as a strictly positive value of $\widehat{\psi}_j$. Even if a zero value of an estimate is an MLE in the sense that the log-likelihood function is maximized at that value, it is not a stationary point of the likelihood equation in the interior of the parameter space. Thus, we restrict ourselves to dealing with the ACM of MLEs of factor loadings under a nearly singular, but not a strictly singular, matrix of MLEs of unique variances in this paper. (See e.g., [5,15] for attempts to include a strictly singular matrix of MLEs of unique variances as long as we are confident in our assumption that Ψ_0 is nonsingular. However, further studies are needed on whether the formulas given below still approximate well the true ACM in such a case.)

Fourth, as the jth diagonal element $\widehat{\psi}_j$ of the matrix of MLEs of unique variances gets closer and closer to zero, the matrix of MLEs of factor loadings follows a specific pattern: the jth row of the matrix of MLEs of factor loadings approaches zero, except for the $(j, 1)$ element, which gets close to the square root of the jth diagonal element of sample covariance matrix (except for the sign change). We state this as the following observation.

Observation 2. *If $\widehat{\psi}_j \approx 0$ (with the rest of the MLEs of unique variances not nearly zero), then $\widehat{\lambda}_{j1} \approx s_{jj}^{1/2}$ (or $\widehat{\lambda}_{j1} \approx -s_{jj}^{1/2}$) and $\widehat{\lambda}_{jr} \approx 0,\ r = 2, \ldots, m$.*

The proof is given in Appendix A.2. This result must have been known by Jöreskog [14], who described the phenomenon (e.g., in his Table 5) and developed a partialing procedure to yield estimates of $\lambda_{jr} = \epsilon_r$ with the exact property that $\epsilon_r = 0$.

4. The ACM of MLEs of factor loadings and unique variances: Σ version

We now give alternative formulas for the asymptotic covariance matrix of MLEs of factor loadings and unique variances which can be used even when one element of estimated unique variances is very close to zero. As ψ_j approaches zero, Eqs. (5) and (13) become unstable (because they involve ψ_j^{-1}, which diverges). Thus it is necessary to replace these equations with alternative formulas that do not involve ψ_j^{-1} (see e.g., [4]). Our approach is motivated by theirs. In addition, the $(1, 1)$ element of Θ also gets very large as ψ_j approaches zero. Thus, Eqs. (8)–(12) also need to be modified. Our modified formulas are as follows:

$$\begin{aligned} A = {} & \{M \otimes \Sigma + (M \otimes \Lambda M)(\mathrm{diag}(K_m \gamma^*))(I_m \otimes \Lambda')\} \\ & -\{A_1 \odot A_1' \odot A_2\}, \end{aligned} \tag{17}$$

$$\begin{aligned} B = {} & -\{\Sigma^{-1}\Lambda M \otimes 1_p'\} \\ & \odot\{1_m' \otimes I_p + (1_m' \otimes \Sigma^{-1}\Lambda)(\mathrm{diag}(K_m \nu^b))(I_m \otimes \Lambda')\}, \end{aligned} \tag{18}$$

$$E = (\Phi \odot \Phi)^{-1}, \tag{19}$$

with

$$A_1 = 1_m \otimes \Lambda \otimes 1_p' - (\mathrm{diag}(\lambda))(I_m \otimes 1_p 1_p'), \tag{20}$$

$$A_2 = N N^{\ddagger} N \otimes 1_p 1_p', \tag{21}$$

$$\gamma^* = \mathrm{vec}((I_m - N)^2 N^{\ddagger} N^2 - 1_m 1_m' + (\tfrac{1}{2}) I_m), \tag{22}$$

$$M = (I_m - N)^{-1}, \tag{23}$$

$$N^{\ddagger} = \mathrm{vec}^{-1}(((N \otimes I_m - I_m \otimes N)^2)^+ 1_{m^2}), \tag{24}$$

$$\nu^b = (N \otimes I_m - I_m \otimes N - 2\,\mathrm{diag}(\mathrm{vec}(I_m - N)))^{-1} 1_{m^2}, \tag{25}$$

$$\Phi = \Sigma^{-1} - \Sigma^{-1}\Lambda M \Lambda' \Sigma^{-1}, \tag{26}$$

where N is an $m \times m$ diagonal matrix whose elements are the m smallest eigenvalues (in ascending order) of $U\Psi U'$ with $\Sigma^{-1} = U'U$, where U is an upper triangular matrix with positive diagonal elements obtained by the Cholesky decomposition of Σ^{-1}. (The elementwise formulas corresponding to (17) and (18) are given in Appendix A.4.)

It is easy to show the equivalence between Eqs. (4)–(13) and Eqs. (17)–(26), by noting the identities $N = \Theta^{-1}$, $\Psi^{-1}\Lambda\Theta^{-1} = \Sigma^{-1}\Lambda$, and $\Psi^{-1}\Lambda(\Theta - I_m)^{-1} =$

$\Sigma^{-1}\Lambda M$. See Appendix A.5 for the proof of the equivalence. As before, we assume that the determinant $|\Phi \odot \Phi| \neq 0$ and thus $\Phi \odot \Phi$ is nonsingular, so that we can compute E in (19) using Φ in (26). (Φ in itself is in general not of full rank, but of rank $p - m$, see, e.g., [1, p. 23]).

In conclusion, the ACM of MLEs of factor loadings and unique variances can be computed using Eqs. (3), (17)–(26), in place of Eqs. (3), (4)–(13), including when one element of the MLEs of unique variances is nearly zero.

5. Alternative matrix formulas: Ψ version and Σ version

Alternatively, it is possible to construct the matrix formulas for the ACM of estimates of factor loadings and unique variances based on the differentials reported in [11,13] with some modifications. The Ψ version of formulas is as follows:

$$\begin{pmatrix} \mathrm{Var}(\widehat{\lambda}) & \mathrm{Cov}(\widehat{\lambda}, \widehat{\psi}) \\ \mathrm{Cov}(\widehat{\psi}, \widehat{\lambda}) & \mathrm{Var}(\widehat{\psi}) \end{pmatrix} = \begin{pmatrix} \frac{\partial\lambda}{\partial\sigma'} \\ \frac{\partial\psi}{\partial\sigma'} \end{pmatrix} (\mathrm{Cov}(s)) \left(\left(\frac{\partial\lambda}{\partial\sigma'}\right)' \left(\frac{\partial\psi}{\partial\sigma'}\right)' \right), \tag{27}$$

where the matrices of partial derivatives of $\widehat{\lambda}$ and $\widehat{\psi}$ involving Ψ are given by

$$\frac{\partial\lambda}{\partial\sigma'} = (\Gamma^{-1}\Xi \otimes I_p)\left(G_p - \left(\frac{\partial\psi}{\partial\sigma'}\right)\right) - (\Gamma^{-1} \otimes \Lambda)(\Upsilon_1 + \Upsilon_2), \tag{28}$$

$$\frac{\partial\psi}{\partial\sigma'} = K_p^*(\Phi \odot \Phi)^{-1} K_p^{*\prime}(\Phi \otimes \Phi)G_p, \tag{29}$$

with

$$\Gamma = \Lambda'\Psi^{-1}\Lambda, \quad \Xi = \Lambda'\Psi^{-1}, \quad \Phi = \Psi^{-1} - \Xi'\Gamma^{-1}\Xi,$$

$$K_p^* = \sum_{i=1}^{p}(J_{p,i}J_{p,i}' \otimes J_{p,i}),$$

and

$$\Upsilon_1 = \left(\frac{1}{2}\right)(\mathrm{diag}(\mathrm{vec}(\Gamma^{-1})))(\Xi \otimes \Xi)\left(G_p - \left(\frac{\partial\psi}{\partial\sigma'}\right)\right), \tag{30}$$

$$\Upsilon_2 = (I_m \otimes \Gamma - \Gamma \otimes I_m)^+ \left\{(\Xi \otimes \Xi)G_p - ((\Gamma + I_m)\Xi \otimes \Xi)\left(\frac{\partial\psi}{\partial\sigma'}\right)\right\}, \tag{31}$$

with the p-dimensional ith unit vector $J_{p,i}$, and under normal sampling, $\mathrm{Cov}(s)$ is given by

$$\mathrm{Cov}(s) = (\tfrac{1}{n})H_p(I_{p^2} + K_p)(\Sigma \otimes \Sigma)H_p', \tag{32}$$

where K_p is the $p^2 \times p^2$ commutation matrix (i.e., $K_p\,\mathrm{vec}(A) = \mathrm{vec}(A')$ for any $p \times p$ matrix A), and $H_p = (G_p'G_p)^{-1}G_p'$ and G_p is the $p^2 \times p(p+1)/2$ duplication matrix (i.e., vec(S) = $G_p$vech(S) for any $p \times p$ symmetric matrix S). (Essentially the identical expression to $\partial\psi/\partial\sigma'$ is obtained by Ihara and Kano [8].) The proof for the alternative matrix approach (Ψ version) is given in Appendix A.6.

The Σ version of the formulas (see also [7]) replaces the matrix of partial derivatives of $\widehat{\lambda}$ and $\widehat{\psi}$ in (28) and (29) by

$$\frac{\partial\lambda}{\partial\sigma'} = (W^{-1}Z \otimes I_p)\left(G_p - \left(\frac{\partial\psi}{\partial\sigma'}\right)\right) - (W^{-1} \otimes \Lambda)(Y_1 + Y_2), \tag{33}$$

$$\frac{\partial\psi}{\partial\sigma'} = K_p^*(Q \odot Q)^{-1}K_p^{*\prime}(Q \otimes Q)G_p, \tag{34}$$

respectively, with $W = \Lambda'\Sigma^{-1}\Lambda,\ Z = \Lambda'\Sigma^{-1}$, $Q = I_p - \Lambda W^{-1}Z$, and

$$Y_1 = \left(\frac{1}{2}\right)(\mathrm{diag}(\mathrm{vec}(W^{-1})))(Z \otimes Z)\left(G_p - \left(\frac{\partial\psi}{\partial\sigma'}\right)\right), \tag{35}$$

$$\begin{aligned} Y_2 &= (I_m \otimes W - W \otimes I_m)^+ \\ &\quad \times \left\{((I_m - W)Z \otimes Z)G_p - (Z \otimes Z)\left(\frac{\partial\psi}{\partial\sigma'}\right)\right\}. \end{aligned} \tag{36}$$

See Appendix A.7 for the proof of the partial derivatives in the Σ version.

6. The augmented information matrix approach

An alternative method to obtain the ACM of the MLEs of factor loadings is to consider it as a constrained MLE problem, using the augmented information matrix [9,17]. The augmented information approach gives a procedure to compute the ACM, but it does not give explicit formulas for the elements of the ACM. However, this approach is easy to implement; it is applicable to other rotated solutions as well; therefore it is a very practical approach. In this section, we consider modifying the standard augmented information matrix approach so that it can be used even when an element of the matrix of MLEs of unique variances is nearly zero.

In case of the unrotated, unstandardized factor loadings, the formulas for the elements of the information matrix are given by:

$$\begin{aligned} x_{ir,js} &= \left(-\frac{1}{n}\right)\boldsymbol{E}\left(\frac{\partial^2 L}{\partial\lambda_{ir}\partial\lambda_{js}}\right) \\ &= \sigma^{ij}(\Lambda'\Sigma^{-1}\Lambda)_{rs} + (\Sigma^{-1}\Lambda)_{is}(\Sigma^{-1}\Lambda)_{jr}, \end{aligned} \tag{37}$$

$$y_{ir,j} = \left(-\frac{1}{n}\right)\boldsymbol{E}\left(\frac{\partial^2 L}{\partial\lambda_{ir}\partial\psi_j}\right) = \sigma^{ij}(\Sigma^{-1}\Lambda)_{jr}, \tag{38}$$

$$z_{i,j} = \left(-\frac{1}{n}\right)\boldsymbol{E}\left(\frac{\partial^2 L}{\partial\psi_i\partial\psi_j}\right) = \left(\frac{1}{2}\right)(\sigma^{ij})^2, \tag{39}$$

for $1 \leqslant i, j \leqslant p,\ 1 \leqslant r, s \leqslant m$ [9], where L is the log likelihood function, σ^{ij} the (i, j) element of Σ^{-1}, and $(V)_{rs}$ is the (r, s) element of matrix V. The $m(m-1)/2$ constraint functions g_{uv} on the parameters λ_{ir} and ψ_j are

$$g_{uv} = (\Lambda'\Psi^{-1}\Lambda)_{uv} \tag{40}$$

for $1 \leqslant u < v \leqslant m$, and the partial derivatives of g_{uv} with respect to λ_{ir} and ψ_j are given by

$$f^1_{ir,uv} = \frac{\partial g_{uv}}{\partial \lambda_{ir}} = (\delta_{ru}\lambda_{iv} + \delta_{rv}\lambda_{iu})\psi_i^{-2}, \tag{41}$$

$$f^2_{j,uv} = \frac{\partial g_{uv}}{\partial \psi_j} = -\lambda_{ju}\lambda_{jv}\psi_j^{-2}. \tag{42}$$

Now define the matrices $X = (x_{ir,js}),\ Y = (y_{ir,j})$, and $Z = (z_{i,j})$ from the coordinatewise expressions in Eqs. (37)–(39). (Note that the subscripts r and s serve as row and column block indices, respectively, while i and j are row and column indices within each block. That is, $x_{ir,js}$ is the $((r-1)p+i, (s-1)p+j)$ element of X, $y_{ir,j}$ the $((r-1)p+i, j)$ element of Y, and $z_{i,j}$ is the (i, j) element of Z. The orders of $X,\ Y$, and Z are $pm \times pm,\ pm \times p$, and $p \times p$, respectively. X and Z are symmetric matrices. Likewise, define the matrices of partial derivatives $F_1 = (f^1_{ir,uv})$ and $F_2 = (f^2_{j,uv})$ from the coordinatewise expressions in Eqs. (41) and (42). ($f^1_{ir,uv}$ is the $((r-1)p+i, u(2m-u-1)/2+v-m)$ element of F_1, $f^2_{j,uv}$ is the $(j, u(2m-u-1)/2+v-m)$ element of F_2. The orders of F_1 and F_2 are $pm \times m(m-1)/2$ and $p \times m(m-1)/2$, respectively.) Then the augmented information matrix is given by the sample size n times the $(p(m+1)+m(m-1)/2) \times (p(m+1)+m(m-1)/2)$ matrix whose submatrices are arranged as

$$\begin{pmatrix} X & Y & F_1 \\ Y' & Z & F_2 \\ F_1' & F_2' & 0 \end{pmatrix}, \tag{43}$$

and the ACM for the MLEs of factor loadings and unique variances is the $p(m+1) \times p(m+1)$ submatrix corresponding to the first $p(m+1)$ rows and columns of the inverse of the augmented information matrix.

Here, note that $x_{ir,js}$, $y_{ir,j}$, and $z_{i,j}$ in (37)–(39) are functions of the elements of Σ^{-1} and Λ, which are all finite. Thus the functions $x_{ir,js},\ y_{ir,j}$, and $z_{i,j}$ are all finite, and the estimates of $x_{ir,js},\ y_{ir,j}$, and $z_{i,j}$ can be computed without any modification of the formulas even when the MLE of ψ_j is nearly zero. On the other hand, the equations for the partial derivatives $f^1_{ir,uv}$ and $f^2_{j,uv}$ in (41) and (42) involve ψ_j^{-1} and thus for some elements, the estimates of $f^1_{ir,uv}$ and $f^2_{j,uv}$ become very large when the MLE of ψ_j is nearly zero.

Motivated by Bentler and Yuan [4], Okamoto [19], and Swain [20], we use the alternative constraint functions $h_{uv} = (\Lambda'\Sigma^{-1}\Lambda)_{uv}$, instead of $g_{uv} = (\Lambda'\Psi^{-1}\Lambda)_{uv}$ in (30), when an element of MLE of Ψ is nearly zero. In fact, the constraint that $\Lambda'\Psi^{-1}\Lambda$ is diagonal is equivalent to the constraint that $\Lambda'\Sigma^{-1}\Lambda$ is diagonal, except that the former constraint requires the assumption that Ψ is positive definite, while the latter constraint requires the assumption that Σ is positive definite. To see the equivalence of the two constraints in the regular case, first note the identity:

$$\Lambda'\Sigma^{-1}\Lambda = \Lambda'\Psi^{-1}\Lambda(I_m + \Lambda'\Psi^{-1}\Lambda)^{-1} \tag{44}$$

and note that, since the RHS of (44) is diagonal, the LHS of (44) also has to be diagonal.

The partial derivatives of h_{uv} with respect to λ_{ir} and ψ_j, which are used inside the augmented information matrix are given by:

$$\frac{\partial h_{uv}}{\partial \lambda_{ir}} = (J_{ir}'\Sigma^{-1}\Lambda + \Lambda'\Sigma^{-1}J_{ir} - \Lambda'\Sigma^{-1}(J_{ir}\Lambda' + \Lambda J_{ir}')\Sigma^{-1}\Lambda)_{uv}, \tag{45}$$

$$\frac{\partial h_{uv}}{\partial \psi_j} = -(\Lambda'\Sigma^{-1}K_{jj}\Sigma^{-1}\Lambda)_{uv}, \tag{46}$$

where J_{ir} is a $p \times m$ matrix whose (i, r) element is 1 and the rest of the elements are all zero, and K_{jj} is a $p \times p$ matrix whose (j, j) element is 1 and the rest of the elements are all zero (see [9, p. 125]). The proof for (45) and (46) is given in Appendix A.9.

Thus to compute the ACM of MLEs of factor loadings and unique variances when one element of the MLEs of unique variances is nearly zero, we recommend using the partial derivatives of h_{uv} with respect to λ_{ir} and ψ_j given in (45) and (46), in place of the partial derivatives of g_{uv} with respect to λ_{ir} and ψ_j in (41) and (42). The ACM of MLEs of factor loadings and unique variances is given by the $p(m+1) \times p(m+1)$ submatrix corresponding to the first $p(m+1)$ rows and columns of the inverse of the augmented information matrix with F_1 and F_2 replaced by the partial derivatives of h_{uv}. We should note that the augmented information matrix approach with the partial derivatives of the alternative constraint functions can be used whether or not an element of the MLEs of unique variances is nearly zero.

7. Conclusion

In this paper, we dealt with the ACM of MLEs of (unstandardized, unrotated) factor loadings and unique variances when an element of MLEs of unique variances is nearly zero, that is, the matrix of MLEs of unique variances is nearly singular. The standard formulas for the ACM given by Lawley and Maxwell [18] involve the inverse of the unique variances. Thus, we encounter a problem when one of the MLEs of unique variances approaches zero, since the reciprocal of the MLE of this unique variance gets very large.

We presented alternative formulas for the ACM which can be used even when an element of MLEs of unique variances is nearly zero. The derivation of the alternative formulas involved replacing the expressions in terms of the inverse of unique variances by the expressions in terms of the inverse of the covariance matrix Σ whose elements are all finite. The alternative formulas given in Sections 4 and 5 are exact asymptotic formulas, and they can be used whether or not an element of MLEs of unique variances is nearly zero. In this regard, we consider use of the alternative formulas to be more practical than the standard formulas.

However, it should be noted that statistical instabilities that might arise in the case of unique variances near zero cannot be avoided just by using alternative formulas. For example, it is just possible that the asymptotic standard error of the MLE of a unique variance that is near zero is a very bad approximation to the true standard error in small samples. (The authors thank the referee for noting this important point.) This is the area where we certainly need further research. See also [5,15] on the issues closely related with this point.

Furthermore, we used alternative constraint functions $h_{uv} = (\Lambda'\Sigma^{-1}\Lambda)_{uv}$ instead of $g_{uv} = (\Lambda'\Psi^{-1}\Lambda)_{uv}$, in the context of the augmented information matrix approach, when an element of the MLE of unique variances is nearly zero. In fact, use of the constraint that $\Lambda'\Sigma^{-1}\Lambda$ is diagonal is not new; for example, it was mentioned by Swain [20]. However, to our knowledge, use of the constraint function $h_{uv} = (\Lambda'\Sigma^{-1}\Lambda)_{uv}$ in the context of the augmented information matrix approach for avoiding use of the inverse of unique variances, as well as the formulas for the partial derivatives of h_{uv} with respect to λ_{ir} and ψ_j to be used inside the augmented information matrix, are new.

The augmented information approach has an advantage in that this approach is applicable to other rotated solutions as well. In fact, only the matrices F_1 and F_2 of the partial derivatives of the constraint functions need to be modified for obtaining the standard errors for various rotated solutions. As long as the formulas for the constraint functions are not very complex, in general, the partial derivatives can be obtained fairly easily.

Appendix A

A.1. Proof of Observation 1

(i) By definition of the eigenvalue–eigenvector equation, $(U\Psi U')Z^* = Z^*N^*$. Rearrange this equation to:

$$(\Psi^{-1/2}S\Psi^{-1/2})(\Psi^{1/2}U'Z^*N^{*-1/2}) = (\Psi^{1/2}U'Z^*N^{*-1/2})N^{*-1}, \tag{A.1}$$

and comparing (A.1) with the eigenvalue–eigenvector equation $(\Psi^{-1/2}S\Psi^{-1/2})\Omega^* = \Omega^*\Theta^*$ gives (14) and (15).

(ii) Use Eq. (4.5) of Lawley and Maxwell [18]: $(S - \Sigma)\Sigma^{-1}\Lambda = 0$, and rearrange it as

$$(U\Psi U')(U\Lambda(\Lambda'S^{-1}\Lambda)^{-1/2}) = (U\Lambda(\Lambda'S^{-1}\Lambda)^{-1/2})(I_m - \Lambda'S^{-1}\Lambda). \tag{A.2}$$

(See, e.g., [12,19]). Comparing (A.2) with the eigenvalue–eigenvector equation $(U\Psi U')Z^* = Z^*N^*$ gives $N^* = I_m - \Lambda' S^{-1}\Lambda$ and $Z^* = U\Lambda(\Lambda' S^{-1}\Lambda)^{-1/2} = U\Lambda(I_m - N^*)^{-1/2}$. Thus (16) follows.

A.2. *Proof of Observation 2*

We omit the hat notation thereafter for simplicity. The rth largest eigenvalues $\theta_r,\ r = 2,\ldots,m$, of $\Psi^{-1/2}S\Psi^{-1/2}$, except the largest eigenvalue, do not get very large as long as ψ_j is the only unique variance which is nearly zero. The (j,r) element of $\Lambda = \Psi^{1/2}\Omega^*(\Theta^* - I_m)^{1/2}$ in (1) is $\lambda_{jr} = \psi_j^{1/2}\omega^*_{jr}(\theta^*_r - 1)^{1/2}$. Let ϵ be a positive quantity very close to zero. When $r \neq 1$, $\psi_j = \epsilon$ with bounded ω_{jr} ($|\omega_{jr}| \leqslant 1$) and finite (not very large) θ_r gives $\lambda_{jr} = \epsilon_r,\ r = 2,\ldots,m$, where ϵ_r are quantities very close to zero. Thus, the (j,j) element of $S \approx \Lambda\Lambda' + \Psi$ is $s_{jj} \approx \lambda_{j1}^2 + (\epsilon_2^2 + \cdots + \epsilon_m^2 + \epsilon)$, and $\lambda_{j1} \approx s_{jj}^{1/2}$ or $\lambda_{j1} \approx -s_{jj}^{1/2}$ follows.

A.3. *Outline of proof of the equivalence of the standard elementwise expressions and the matrix expressions (4), (5), (7)–(12)*

For $i, j = 1,\ldots,p$, and $r, s = 1,\ldots,m$, the standard elementwise formulas for A and B (Ψ version: [13,18]) are

$$a_{ir,jr} = \mu_r\left\{\sigma_{ij} - \left(\frac{1}{2}\right)\mu_r\lambda_{ir}\lambda_{jr} + \sum_{k\neq r}^{m}(\mu_k\gamma_{rk}\lambda_{ik}\lambda_{jk})\right\}, \tag{A.3}$$

$$a_{ir,js} = -\{\theta_r\theta_s(\theta_r - \theta_s)^{-2}\}\lambda_{is}\lambda_{jr} \quad \text{for } r \neq s, \tag{A.4}$$

$$\begin{aligned} b_{j,ir} = {} & -\lambda_{jr}(\theta_r - 1)^{-1}\psi_j^{-2} \\ & \times\left\{\delta_{ij}\psi_j - \left(\frac{1}{2}\right)\lambda_{ir}\lambda_{jr}(\theta_r - 1)^{-1} + \theta_r\sum_{k\neq r}^{m}(\lambda_{ik}\lambda_{jk}(\theta_r - \theta_k)^{-1})\right\}, \end{aligned} \tag{A.5}$$

where θ_r is the rth element of the $m \times m$ diagonal matrix Θ which has as its elements the m largest eigenvalues of $\Psi^{-1/2}\Sigma\Psi^{-1/2}$; $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$; and μ_r and γ_{rk} are defined in terms of θ_r's as follows:

$$\mu_r = \frac{\theta_r}{\theta_r - 1}, \tag{A.6}$$

$$\gamma_{rk} = \left(\frac{\theta_r - 1}{\theta_r - \theta_k}\right)^2 - 1. \tag{A.7}$$

First, we show that (A.3) is an element of the first term in (4). Express (A.3) as $a_{ir,jr} = \mu_r(\sigma_{ij} + \sum_{k=r}^{m}(\mu_k\gamma^*_{rk}\lambda_{ik}\lambda_{jk}))$, where $\gamma^*_{rk} = \gamma_{rk}$ if $k \neq r$, $\gamma^*_{rk} = -1/2$ if

$k = r$, and write the first term in (4) as $M \otimes \Sigma + (M \otimes \Lambda M)(\mathrm{diag}(K_m\gamma^*))(I_m \otimes \Lambda') = (M \otimes I_p)\{I_m \otimes \Sigma + (I_m \otimes \Lambda)(I_m \otimes M)(\mathrm{diag}(\mathrm{vec}(\Gamma^{*\prime})))(I_m \otimes \Lambda')\}$ with $\Gamma^* = \mathrm{vec}^{-1}(\gamma^*)$. The equivalence of Γ^* and γ^*_{rk} is obvious by comparing $(\Theta - I_m)^2$ and $(\theta_r - 1)^2$, and $I_m \otimes \Theta - \Theta \otimes I_m$ and $\theta_r - \theta_k$. The rest is straightforward by noting that $M \otimes I_p$ in (4) corresponds to $\mu_r, I_m \otimes \Sigma$ corresponds to σ_{ij}, and $(I_m \otimes \Lambda)(I_m \otimes M)(\mathrm{diag}(\mathrm{vec}(\Gamma^{*\prime})))(I_m \otimes \Lambda')$ corresponds to $\sum_{k=1}^{m}(\mu_k\gamma^*_{rk}\lambda_{ik}\lambda_{jk})$.

Next, we show that (A.4) is an element of the second term in (4). First, notice that A_1 in (7) corresponds to λ_{is} in (A.4). $1_m \otimes \Lambda \otimes 1'_p$ includes both the cases of $r = s$ and $r \neq s$. To set the block diagonal elements which correspond to the case of $r = s$ equal to zero, we need to subtract a block diagonal matrix $(\mathrm{diag}(\lambda))(I_m \otimes 1_p 1'_p)$ from $1_m \otimes \Lambda \otimes 1'_p$. Similarly, A'_1 in (7) corresponds to λ_{jr} in (A.4). Next, A_2 in (8) corresponds to $\theta_r\theta_s/(\theta_r - \theta_s)^2$ in $a_{ir,js}$ in (A.4) and it is easy to see that $\Theta^{\ddagger}$ in (11) corresponds to $(\theta_r - \theta_s)^{-2}$.

It is obvious that the first term in B in (5) corresponds to $\lambda_{jr}(\theta_r - 1)^{-1}\psi_j^{-2}$ in $b_{j,ir}$ in (A.5). Express

$$-\left(\frac{1}{2}\right)\lambda_{ir}\lambda_{jr}(\theta_r - 1)^{-1} + \theta_r\sum_{k\neq r}^{m}(\lambda_{ik}\lambda_{jk}(\theta_r - \theta_k)^{-1})$$

as $\theta_r\sum_{k=1}^{m}\lambda_{ik}\lambda_{jk}\theta^b_{rk}$, where $\theta^b_{rk} = (\theta_r - \theta_k)^{-1}$ if $r \neq k$ and $\theta^b_{rk} = -(2\theta_r(\theta_r - 1))^{-1}$ if $r = k$. The equivalence of θ^b and θ^b_{rk} is obvious by noting that $I_m \otimes \Theta - \Theta \otimes I_m$ corresponds to $\theta_r - \theta_k$ and $\Theta(\Theta - I_m)$ corresponds to $\theta_r(\theta_r - 1)$. The rest is straightforward by noting that $(1'_m\Theta \otimes \Lambda)(\mathrm{diag}(K_{mm}\theta^b))(I_m \otimes \Lambda')$ corresponds to $\theta_r\sum_{k=1}^{m}(\lambda_{ik}\lambda_{jk}\theta^b_{rk})$.

A.4. Alternative elementwise formulas for A and B (Σ version)

For $i, j = 1, \ldots, p$, and $r, s = 1, \ldots, m$,

$$a_{ir,jr} = \mu_r\left\{\sigma_{ij} - \left(\frac{1}{2}\right)\mu_r\lambda_{ir}\lambda_{jr} + \sum_{k\neq r}^{m}(\mu_k\gamma_{rk}\lambda_{ik}\lambda_{jk})\right\}, \tag{A.8}$$

$$a_{ir,js} = -\{\nu_r\nu_s(\nu_s - \nu_r)^{-2}\}\lambda_{is}\lambda_{jr} \quad \text{for } r \neq s, \tag{A.9}$$

$$\begin{aligned} b_{j,ir} = -\mu_r\left(\sum_{s=1}^{p}\sigma^{js}\lambda_{sr}\right)\Bigg\{\delta_{ij} &- \left(\frac{1}{2}\right)\lambda_{ir}\mu_r\left(\sum_{s=1}^{p}\sigma^{js}\lambda_{sr}\right) \\ &+ \sum_{k\neq r}^{m}\lambda_{ik}(\nu_k - \nu_r)^{-1}\left(\sum_{s=1}^{p}\sigma^{js}\lambda_{sk}\right)\Bigg\}, \end{aligned} \tag{A.10}$$

where ν_r are the eigenvalues of $U\Psi U'$, and μ_r and γ_{rk} are defined in terms of ν_r's as follows:

$$\mu_r = (1 - \nu_r)^{-1}, \tag{A.11}$$

$$\gamma_{rk} = \nu_k^2 \left(\frac{1-\nu_r}{\nu_k - \nu_r} \right)^2 - 1. \tag{A.12}$$

A.5. Proof of the equivalence of the matrix expressions (Ψ version) and the alternative matrix expressions (Σ version)

(i) To show the equivalence of (5) and (18):

$$\begin{aligned}
B &= -\{\Psi^{-2}\Lambda(\Theta - I_m)^{-1} \otimes 1_p'\} \\
&\quad \odot\{1_m' \otimes \Psi + (1_m'\Theta \otimes \Lambda)(\mathrm{diag}(K_m\theta^b))(I_m \otimes \Lambda')\} \\
&= -\{\Psi^{-1}\Lambda(\Theta - I_m)^{-1} \otimes 1_p'\} \\
&\quad \odot\{1_m' \otimes I_m + (1_m'\Theta \otimes \Psi^{-1}\Lambda)(\mathrm{diag}(K_m\theta^b))(I_m \otimes \Lambda')\} \\
&= -\{\Sigma^{-1}\Lambda M \otimes 1_p'\} \\
&\quad \odot\{1_m' \otimes I_m + (1_m'\Theta \otimes \Psi^{-1}\Lambda)(\mathrm{diag}(K_m\theta^b))(I_m \otimes \Lambda')\} \\
&= -\{\Sigma^{-1}\Lambda M \otimes 1_p'\} \\
&\quad \odot\{1_m' \otimes I_m + (1_m' \otimes \Psi^{-1}\Lambda\Theta^{-1})(\Theta \otimes \Theta)(\mathrm{diag}(K_m\theta^b))(I_m \otimes \Lambda')\} \\
&= -\{\Sigma^{-1}\Lambda M \otimes 1_p'\} \\
&\quad \odot\{1_m' \otimes I_m + (1_m' \otimes \Sigma^{-1}\Lambda)\{\mathrm{diag}((\Theta \otimes \Theta)K_m\theta^b)\}(I_m \otimes \Lambda')\} \\
&= -\{\Sigma^{-1}\Lambda M \otimes 1_p'\} \\
&\quad \odot\{1_m' \otimes I_m + (1_m' \otimes \Sigma^{-1}\Lambda)\{\mathrm{diag}(K_m(\Theta \otimes \Theta)\theta^b)\}(I_m \otimes \Lambda')\} \\
&= -\{\Sigma^{-1}\Lambda M \otimes 1_p'\} \\
&\quad \odot\{1_m' \otimes I_p + (1_m' \otimes \Sigma^{-1}\Lambda)(\mathrm{diag}(K_m\nu^b))(I_m \otimes \Lambda')\},
\end{aligned}$$

since

$$\Psi^{-1}\Lambda\Theta^{-1} = \Sigma^{-1}\Lambda, \quad \Psi^{-1}\Lambda(\Theta - I_m)^{-1} = \Sigma^{-1}\Lambda M,$$

and

$$\begin{aligned}
&(\Theta \otimes \Theta)\theta^b \\
&\quad = (\Theta \otimes \Theta)(I_m \otimes \Theta - \Theta \otimes I_m - 2\,\mathrm{diag}(\mathrm{vec}(\Theta(\Theta - I_m))))^{-1}1_{m^2} \\
&\quad = (\Theta^{-1} \otimes I_m - I_m \otimes \Theta^{-1} - 2(\Theta^{-1}\otimes\Theta^{-1})\mathrm{diag}(\mathrm{vec}(\Theta(\Theta - I_m))))^{-1}1_{m^2} \\
&\quad = [\Theta^{-1} \otimes I_m - I_m \otimes \Theta^{-1} - 2\mathrm{diag}\{\mathrm{vec}((\Theta - I_m)\Theta^{-1})\}]^{-1}1_{m^2} \\
&\quad = (N \otimes I_m - I_m \otimes N - 2\mathrm{diag}(\mathrm{vec}(I_m - N)))^{-1}1_{m^2} \\
&\quad = \nu^b.
\end{aligned}$$

(ii) To show the equivalence of (8) and (21):

$$
\begin{aligned}
&\mathrm{vec}(\Theta\Theta^{\ddagger}\Theta)\\
&\quad= (\Theta\otimes\Theta)\mathrm{vec}\Theta^{\ddagger} = (\Theta\otimes\Theta)((I_m\otimes\Theta-\Theta\otimes I_m)^2)^{+}1_{m^2}\\
&\quad= ((\Theta^{-1}\otimes\Theta^{-1})(I_m\otimes\Theta-\Theta\otimes I_m)^2)^{+}1_{m^2}\\
&\quad= ((\Theta\otimes\Theta)(\Theta^{-1}\otimes\Theta^{-1})^2(I_m\otimes\Theta-\Theta\otimes I_m)^2)^{+}1_{m^2}\\
&\quad= (\Theta^{-1}\otimes\Theta^{-1})((\Theta^{-1}\otimes I_m-I_m\otimes\Theta^{-1})^2)^{+}1_{m^2}\\
&\quad= (N\otimes N)\{((N\otimes I_m-I_m\otimes N)^2)^{+}1_{m^2}\}\\
&\quad= (N\otimes N)\mathrm{vec}N^{\ddagger}\\
&\quad= \mathrm{vec}(NN^{\ddagger}N).
\end{aligned}
$$

(iii) To show the equivalence of (9) and (22):

$$
\begin{aligned}
&\mathrm{vec}((\Theta-I_m)^2\Theta^{\ddagger})\\
&\quad= (I_m\otimes(\Theta-I_m)^2)((I_m\otimes\Theta-\Theta\otimes I_m)^2)^{+}1_{m^2}\\
&\quad= \{\{(I_m\otimes(\Theta-I_m))^{-1}(I_m\otimes\Theta-\Theta\otimes I_m)\}^2\}^{+}1_{m^2}\\
&\quad= \{\{(I_m\otimes(\Theta-I_m)^{-1})(I_m\otimes\Theta-\Theta\otimes I_m)\}^2\}^{+}1_{m^2}\\
&\quad= \{\{I_m\otimes(\Theta-I_m)^{-1}\Theta-\Theta\otimes(\Theta-I_m)^{-1}\}^2\}^{+}1_{m^2}\\
&\quad= \{\{I_m\otimes(I_m-\Theta^{-1})^{-1}-\Theta\otimes(I_m-\Theta^{-1})^{-1}\Theta^{-1}\}^2\}^{+}1_{m^2}\\
&\quad= \{\{I_m\otimes(I_m-N)^{-1}-N^{-1}\otimes(I_m-N)^{-1}N\}^2\}^{+}1_{m^2}\\
&\quad= \{(N^{-2}\otimes(I_m-N)^{-2})\{N\otimes I_m-I_m\otimes N\}^2\}^{+}1_{m^2}\\
&\quad= (N^2\otimes(I_m-N)^2)\mathrm{vec}N^{\ddagger}\\
&\quad= \mathrm{vec}((I_m-N)^2N^{\ddagger}N^2).
\end{aligned}
$$

(iv) To show the equivalence of (10) and (23): Using $N=\Theta^{-1}$,

$$M=\Theta(\Theta-I_m)^{-1}=N^{-1}(N^{-1}-I_m)^{-1}=(NN^{-1}-N)^{-1}=(I_m-N)^{-1}.$$

(v) To show the equivalence of (13) and (26): Use the identity $\Sigma^{-1}=\Psi^{-1}-\Psi^{-1}\Lambda\Theta^{-1}\Lambda'\Psi^{-1}$, and subtract this from Φ in (13):

$$
\begin{aligned}
\Phi-\Sigma^{-1}&=\Psi^{-1}\Lambda(\Theta^{-1}-(\Theta-I_m)^{-1})\Lambda'\Psi^{-1}\\
&=\Sigma^{-1}\Lambda\Theta(\Theta^{-1}-(\Theta-I_m)^{-1})\Theta\Lambda'\Sigma^{-1}\\
&=-\Sigma^{-1}\Lambda(\Theta-I_m)^{-1}\Theta\Lambda'\Sigma^{-1}\\
&=-\Sigma^{-1}\Lambda M\Lambda'\Sigma^{-1},
\end{aligned}
$$

since

$$\Psi^{-1}\Lambda\Theta^{-1}=\Sigma^{-1}\Lambda$$

and

$$\Theta^{-1} - (\Theta - I_m)^{-1} = -\Theta^{-1}(\Theta - I_m)^{-1}.$$

A.6. Proof of the expressions for the matrices of partial derivatives of $\widehat{\lambda}$ and $\widehat{\psi}$ (Ψ version)

The likelihood equations are written in the form: $(S - \widehat{\Lambda}\widehat{\Lambda}' - \widehat{\Psi})\widehat{\Psi}^{-1}\widehat{\Lambda} = 0$, $\mathrm{diag}(S - \widehat{\Lambda}\widehat{\Lambda}' - \widehat{\Psi}) = 0$, and $\mathrm{nondiag}(\widehat{\Lambda}'\widehat{\Psi}^{-1}\widehat{\Lambda}) = 0$, the differentials of which are

$$(\mathrm{d}\Sigma - \mathrm{d}\Lambda\Lambda' - \Lambda\,\mathrm{d}\Lambda' - \mathrm{d}\Psi)\Psi^{-1}\Lambda = 0,$$
$$\mathrm{diag}(\mathrm{d}\Sigma - \mathrm{d}\Lambda\Lambda' - \Lambda\,\mathrm{d}\Lambda' - \mathrm{d}\Psi) = 0,$$

and

$$\mathrm{nondiag}(\mathrm{d}\Lambda'\Psi^{-1}\Lambda - \Lambda'\Psi^{-1}\,\mathrm{d}\Psi\Psi^{-1}\Lambda + \Lambda'\Psi^{-1}\,\mathrm{d}\Lambda) = 0,$$

respectively, that is,

$$\mathrm{d}\Lambda\Gamma = (\mathrm{d}\Sigma - \mathrm{d}\Psi)\Xi' - \Lambda\,\mathrm{d}\Lambda'\Xi', \tag{A.13}$$

$$\mathrm{diag}(\mathrm{d}\Sigma - \mathrm{d}\Lambda\Lambda' - \Lambda\,\mathrm{d}\Lambda' - \mathrm{d}\Psi) = 0, \tag{A.14}$$

$$\mathrm{nondiag}(\mathrm{d}\Lambda'\Xi' + \Xi\,\mathrm{d}\Lambda - \Xi\,\mathrm{d}\Psi\Xi') = 0. \tag{A.15}$$

Vectorizing both sides of (A.13) and letting, $\mathrm{d}\lambda^b = \mathrm{vec}(\mathrm{d}\Lambda'\Xi')$, $(\Gamma \otimes I_p)(\mathrm{d}\lambda) = (\Lambda'\Psi^{-1} \otimes I_p)\mathrm{vec}(\mathrm{d}\Sigma - \mathrm{d}\Psi) - (I_m \otimes \Lambda)(\mathrm{d}\lambda^b)$. Thus, letting $\partial\lambda^b/\partial\sigma' = \Upsilon_1 + \Upsilon_2$, where Υ_1 and Υ_2 are the diagonal and the off-diagonal components of $\partial\lambda^b/\partial\sigma'$, respectively, (28) follows.

Now, premultiplying (A.13) by Ξ gives $\Xi\,\mathrm{d}\Lambda\Gamma + \Gamma\mathrm{d}\Lambda'\Xi' = \Xi(\mathrm{d}\Sigma - \mathrm{d}\Psi)\Xi'$. For the diagonal elements of $\mathrm{d}\Lambda'\Xi'$, $2\Gamma\,\mathrm{d}\Lambda'\Xi' = \Xi(\mathrm{d}\Sigma - \mathrm{d}\Psi)\Xi'$, that is,

$$\mathrm{d}\Lambda'\Xi' = \left(\frac{1}{2}\right)\Gamma^{-1} \odot \Xi(\mathrm{d}\Sigma - \mathrm{d}\Psi)\Xi'. \tag{A.16}$$

Vectorizing (A.16) gives

$$\begin{aligned}
\mathrm{d}\lambda^b &= \left(\frac{1}{2}\right)\mathrm{vec}(\Gamma^{-1}) \odot \mathrm{vec}(\Xi(\mathrm{d}\Sigma - \mathrm{d}\Psi)\Xi') \\
&= \left(\frac{1}{2}\right)\mathrm{vec}(\Gamma^{-1}) \odot (\Xi \otimes \Xi)\mathrm{vec}(\mathrm{d}\Sigma - \mathrm{d}\Psi) \\
&= \left(\frac{1}{2}\right)\mathrm{diag}(\mathrm{vec}(\Gamma^{-1}))(\Xi \otimes \Xi)\mathrm{vec}(\mathrm{d}\Sigma - \mathrm{d}\Psi).
\end{aligned}$$

Thus Υ_1 in (30) follows.

Next, for the nondiagonal elements of $\mathrm{d}\Lambda'\Xi'$, inserting $\mathrm{nondiag}(\Xi\,\mathrm{d}\Lambda) = \mathrm{nondiag}(\Xi\,\mathrm{d}\Psi\Xi' - \mathrm{d}\Lambda'\Xi')$ in (A.15) into $\mathrm{nondiag}(\Xi\,\mathrm{d}\Lambda\Gamma + \Gamma\,\mathrm{d}\Lambda'\Xi') = \mathrm{nondiag}(\Xi(\mathrm{d}\Sigma - \mathrm{d}\Psi)\Xi')$ in (A.14) gives

$$\mathrm{nondiag}(\Gamma\,\mathrm{d}\Lambda'\Xi' - \mathrm{d}\Lambda'\Xi'\Gamma) = \mathrm{nondiag}(\Xi\,\mathrm{d}\Sigma\Xi' - \Xi\,\mathrm{d}\Psi\Xi'(\Gamma + I_m)). \quad \text{(A.17)}$$

The vectorized version of (A.17) is

$$(I_m \otimes \Gamma - \Gamma \otimes I_m)(\mathrm{d}\lambda^b) = (\Xi \otimes \Xi)\mathrm{vec}(\mathrm{d}\Sigma) - ((\Gamma + I_m)\Xi \otimes \Xi)\mathrm{vec}(\mathrm{d}\Psi),$$

from which Υ_2 in (31) follows. Finally, noting $\Phi\Lambda = 0$ and $\Lambda'\Phi = 0$, (A.14) leads to $\mathrm{diag}(\Phi\,\mathrm{d}S\Phi) = \mathrm{diag}(\Phi\,\mathrm{d}\Psi\Phi)$. Then

$$\mathrm{vdg}(\mathrm{d}\Psi) = (\Phi \odot \Phi)^{-1}\mathrm{vdg}(\Phi\mathrm{d}\Sigma\Phi). \quad \text{(A.18)}$$

Vectorizing the diagonal matrix whose elements are (A.18) gives

$$\begin{aligned}
\mathrm{d}\psi &= \mathrm{vec}(\mathrm{d}\Psi) = \mathrm{vec}(\mathrm{diag}((\Phi \odot \Phi)^{-1}\mathrm{vdg}(\Phi(\mathrm{d}\Sigma)\Phi)))\\
&= K_p^*(\Phi \odot \Phi)^{-1}K_p^{*\prime}\mathrm{vec}(\Phi(\mathrm{d}\Sigma)\Phi)\\
&= K_p^*(\Phi \odot \Phi)^{-1}K_p^{*\prime}(\Phi \otimes \Phi)G_p(\mathrm{d}\sigma),
\end{aligned}$$

from which (29) follows.

A.7. Proof of the expressions for the matrices of partial derivatives of $\widehat{\lambda}$ and $\widehat{\psi}$ (Σ version)

The likelihood equations in the Σ version are

$$\begin{aligned}
&(S - \widehat{\Lambda}\widehat{\Lambda}' - \widehat{\Psi})S^{-1}\widehat{\Lambda} = 0,\\
&\mathrm{diag}(S - \widehat{\Lambda}\widehat{\Lambda}' - \widehat{\Psi}) = 0,\\
&\mathrm{nondiag}(\widehat{\Lambda}'S^{-1}\widehat{\Lambda}) = 0,
\end{aligned}$$

and the differentials are

$$(\mathrm{d}\Sigma - \mathrm{d}\Lambda\Lambda' - \Lambda\mathrm{d}\Lambda' - \mathrm{d}\Psi)\Sigma^{-1}\Lambda = 0,$$

(A.14), and

$$\mathrm{nondiag}(\mathrm{d}\Lambda'\Sigma^{-1}\Lambda - \Lambda'\Sigma^{-1}\,\mathrm{d}\Sigma\Sigma^{-1}\Lambda + \Lambda'\Sigma^{-1}\,\mathrm{d}\Lambda) = 0,$$

that is, (A.14) and

$$\mathrm{d}\Lambda W = (\mathrm{d}\Sigma - \mathrm{d}\Psi)Z' - \Lambda\,\mathrm{d}\Lambda'Z', \quad \text{(A.19)}$$

$$\mathrm{nondiag}(\mathrm{d}\Lambda'Z' + Z\,\mathrm{d}\Lambda - Z\,\mathrm{d}\Sigma Z') = 0. \quad \text{(A.20)}$$

Let $\mathrm{d}\lambda^* = \mathrm{d}\Lambda'Z'$. Then vectorizing both sides of (A.19) leads to

$$(W \otimes I_p)(\mathrm{d}\lambda) = (Z \otimes I_p)\mathrm{vec}(\mathrm{d}\Sigma - \mathrm{d}\Psi) - (I_m \otimes \Lambda)(\mathrm{d}\lambda^*),$$

and (33) follows, where Y_1 and Y_2 are the diagonal and off-diagonal components of $\partial\lambda^*/\partial\sigma'$, respectively. For the diagonal components of $\mathrm{d}\Lambda'Z'$, premultiplication of (A.19) by Z leads to $2\,\mathrm{diag}W\,\mathrm{d}\Lambda'Z' = \mathrm{diag}(Z(\mathrm{d}\Sigma - \mathrm{d}\Psi)Z')$, that is,

$$\mathrm{d}\Lambda'Z' = (\tfrac{1}{2})W^{-1} \odot Z(\mathrm{d}\Sigma - \mathrm{d}\Psi)Z'. \quad \text{(A.21)}$$

Vectorizing (A.21) gives

$$\mathrm{d}\lambda^* = (\tfrac{1}{2})\{\mathrm{diag}(\mathrm{vec}(W^{-1}))\}(Z \otimes Z)\mathrm{vec}(\mathrm{d}\Sigma - \mathrm{d}\Psi),$$

from which Y_1 in (35) immediately follows.

For the nondiagonal elements of $\mathrm{d}\Lambda' Z'$, inserting $Z\mathrm{d}\Lambda = Z\mathrm{d}\Sigma Z' - \mathrm{d}\Lambda' Z'$ in (A.20) into (A.19) premultiplied by Z gives

$$W\,\mathrm{d}\Lambda' Z' - \mathrm{d}\Lambda' Z' W = Z\,\mathrm{d}\Sigma Z'(I_m - W) - Z\,\mathrm{d}\Psi Z'. \tag{A.22}$$

Vectorizing (A.22) leads to

$$(I_m \otimes W - W \otimes I_m)(\mathrm{d}\lambda^*) = ((I_m - W)Z \otimes Z)G_p(\mathrm{d}\sigma) - (Z \otimes Z)(\mathrm{d}\psi),$$

from which Y_2 in (36) follows. Now, let $Q = I_p - \Lambda W^{-1} Z$. Then noting $Q\Lambda = 0$ and using (A.19), we have $\mathrm{diag}(Q\,\mathrm{d}\Sigma Q') = \mathrm{diag}(Q\,\mathrm{d}\Psi Q')$, that is, $\mathrm{vdg}(\mathrm{d}\Psi) = (Q \odot Q)^{-1}\mathrm{vdg}(Q\,\mathrm{d}\Sigma Q)$. Thus

$$\begin{aligned}
\mathrm{d}\psi &= \mathrm{vec}(\mathrm{d}\Psi) = \mathrm{vec}(\mathrm{diag}((Q \odot Q)^{-1}\mathrm{vdg}(Q(\mathrm{d}\Sigma)Q)))\\
&= K_p^*(Q \odot Q)^{-1} K^{*\prime}\mathrm{vec}(Q(\mathrm{d}\Sigma)Q)\\
&= K_p^*(Q \odot Q)^{-1} K^{*\prime}(Q \otimes Q)G_p(\mathrm{d}\sigma),
\end{aligned}$$

from which (34) results.

A.8. *Proof of the equivalence of the three equations* $(S - \Sigma)\Psi^{-1}\Lambda = 0$, $(S - \Sigma)S^{-1}\Lambda = 0$, *and* $(S - \Sigma)\Sigma^{-1}\Lambda = 0$ *[19, Theorems 4.2 and 4.3]*

We assume the positive definiteness of Ψ, Σ, and S. Then

$$\begin{aligned}
(S - \Sigma)\Psi^{-1}\Lambda = 0 &\Leftrightarrow (S - \Sigma)\Psi^{-1}\Lambda\Theta^{-1} = 0\\
&\Leftrightarrow (S - \Sigma)\Sigma^{-1}\Lambda = 0 \quad (\text{used } \Psi^{-1}\Lambda\Theta^{-1} = \Sigma^{-1}\Lambda)\\
&\Leftrightarrow S\Sigma^{-1}\Lambda = \Lambda\\
&\Leftrightarrow \Lambda = \Sigma S^{-1}\Lambda\\
&\Leftrightarrow (S - \Sigma)S^{-1}\Lambda = 0.
\end{aligned}$$

A.9. *Proof of the expressions for (35) and (36)*

Take the partial derivatives of $h = \Lambda'\Sigma^{-1}\Lambda$ with respect to λ_{ir} and ψ_j:

$$\begin{aligned}
\frac{\partial h}{\partial \lambda_{ir}} &= \left(\frac{\partial \Lambda'}{\partial \lambda_{ir}}\right)\Sigma^{-1}\Lambda + \Lambda'\left(\frac{\partial \Sigma^{-1}}{\partial \lambda_{ir}}\right)\Lambda + \Lambda'\Sigma^{-1}\left(\frac{\partial \Lambda}{\partial \lambda_{ir}}\right)\\
&= J_{ir}'\Sigma^{-1}\Lambda - \Lambda'\Sigma^{-1}\left(\frac{\partial \Sigma}{\partial \lambda_{ir}}\right)\Sigma^{-1}\Lambda + \Lambda'\Sigma^{-1}J_{ir},
\end{aligned} \tag{A.23}$$

$$
\begin{aligned}
\frac{\partial h}{\partial \psi_j} &= \Lambda' \left(\frac{\partial \Sigma^{-1}}{\partial \psi_j} \right) \Lambda \\
&= -\Lambda' \Sigma^{-1} \left(\frac{\partial \Sigma}{\partial \psi_j} \right) \Sigma^{-1} \Lambda.
\end{aligned} \tag{A.24}
$$

Inserting the following partial derivatives of Σ with respect to λ_{ir} and ψ_j:

$$
\begin{aligned}
\frac{\partial \Sigma}{\partial \lambda_{ir}} &= \left(\frac{\partial \Lambda}{\partial \lambda_{ir}} \right) \Lambda' + \Lambda \left(\frac{\partial \Lambda'}{\partial \lambda_{ir}} \right) = J_{ir} \Lambda' + \Lambda J'_{ir}, \\
\frac{\partial \Sigma}{\partial \psi_j} &= K_{jj},
\end{aligned}
$$

into (A.23) and (A.24), the results follow by taking the (u, v) element.

Acknowledgement

The authors thank the referee for very helpful comments.

References

[1] T.W. Anderson, Estimating linear statistical relationships, Ann. Stat. 12 (1984) 1–45.
[2] T.W. Anderson, Y. Amemiya, The asymptotic normal distribution of estimators in factor analysis under general conditions, Ann. Stat. 16 (1988) 759–771.
[3] T.W. Anderson, H. Rubin, Statistical inference in factor analysis, in: Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, vol. 5, 1956, pp. 111–150.
[4] P.M. Bentler, K.-H. Yuan, Optimal conditionally unbiased equivalent factor score estimators, in: M. Berkane (Ed.), Latent Variables with Applications in Causality, Springer, New York, 1997, pp. 259–281.
[5] D.W. Gerbing, J.C. Anderson, Improper solutions in the analysis of covariance structures: their interpretability and a comparison of alternative respecifications, Psychometrika 52 (1987) 99–111.
[6] T.K. Dijkstra, On statistical inference with parameter estimates on the boundary of the parameter space, British J. Math. Statist. Psychol. 45 (1992) 289–309.
[7] K. Hayashi, P.K. Sen, On covariance estimators of factor loadings in factor analysis, J. Multivar. Anal. 66 (1998) 38–45.
[8] M. Ihara, Y. Kano, Asymptotic equivalence of unique variance estimators in marginal and conditional factor analysis models, Stat. Probab. Lett. 14 (1992) 337–341.
[9] R.I. Jennrich, Simplified formulae for standard errors in maximum-likelihood factor analysis, British J. Math. Stat. Psychol. 27 (1974) 122–131.
[10] R.I. Jennrich, A Gauss–Newton algorithm for exploratory factor analysis, Psychometrika 51 (1986) 277–284.
[11] R.I. Jennrich, D.B. Clarkson, A feasible method for standard errors of estimate in maximum likelihood factor analysis, Psychometrika 45 (1980) 237–247.
[12] R.I. Jennrich, S.M. Robinson, A Newton–Raphson algorithm for maximum likelihood factor analysis, Psychometrika 34 (1969) 111–123.

[13] R.I. Jennrich, D.T. Thayer, A note on Lawley's formulas for standard errors in maximum likelihood factor analysis, Psychometrika 38 (1973) 571–580.
[14] K.G. Jöreskog, Some contributions to maximum likelihood factor analysis, Psychometrika 32 (1967) 443–482.
[15] Y. Kano, Causes and treatment of improper solutions: exploratory factor analysis, Bulletin, vol. 24, The Faculty of Human Sciences, Osaka University, 1998 (in Japanese).
[16] D.N. Lawley, Some new results in maximum likelihood factor analysis, in: Proceedings of the Royal Society of Edinburgh, Section A, vol. 67, 1967, pp. 256–264.
[17] D.N. Lawley, The inversion of an augmented information matrix occurring in factor analysis, in: Proceedings of the Royal Society of Edinburgh, Section A, 1976, pp. 171–178.
[18] D.N. Lawley, A.E. Maxwell, Factor Analysis as a Statistical Method, second ed., Elsevier, New York, 1971.
[19] M. Okamoto, Foundations of Factor Analysis, Nikka-Giren, Tokyo, 1986 (in Japanese).
[20] A.J. Swain, A class of factor analysis estimation procedures with common asymptotic sampling properties, Psychometrika 40 (1975) 315–335.
[21] O.P. van Driel, On various causes of improper solutions in maximum likelihood factor analysis, Psychometrika 43 (1978) 225–243.
[22] H. Yanai, K. Shigemasu, S. Maekawa, M. Ichikawa, Factor Analysis, Asakura-shoten, Tokyo, 1990 (in Japanese).